FILED UNDER SEAL

IN THE UNITED STATES DISTRICT COURT
FOR THE EASTERN DISTRICT OF TEXAS
SHERMAN DIVISION

THE STATE OF TEXAS, et al.,	
Plaintiffs,	
v.	Civil Action No. 4:20-cv-00957-SDJ
GOOGLE LLC,	
Defendant.	

EXHIBIT 4

**PLAINTIFF STATES' OPENING BRIEF TO THE SPECIAL MASTER
FOR THE MARCH 21, 2024 HEARING**

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024

<u>Via E-mail</u>

Eric Mahr
Julie S. Elmer
Freshfields Bruckhaus Deringer LLP
700 13th Street NW, 10th Floor
Washington, DC 20005
Eric.mahr@freshfields.com
julie.elmer@freshfields.com

R. Paul Yetter
Bryce Callahan
Yetter Coleman LLP
811 Main Street, Suite 4100
Houston, Texas 77002
pyetter@yettercoleman.com
bcallahan@yettercoleman.com

John D. Harkrider
Axinn Veltrop & Harkrider, LLP
114 West 47th Street
New York, NY 10036
jharkrider@axinn.com

Daniel S. Bitton
Axinn Veltrop & Harkrider, LLP
560 Mission St.
San Francisco, CA 94105
dbitton@axinn.com

Re: *The State of Texas, et al. v. Google LLC*, Case No. 4:20-cv-00957-SDJ (E.D. Tex.)

Dear Counsel:

We write concerning several deficiencies in Google's document production. *First*, the States have identified a significant number of relevant "dashboards" and reporting tools (and corresponding reports) that Google uses and/or creates in the ordinary course of its Display Ads business that have not been produced or otherwise made available to the States. *Second*, the States have identified a number of relevant source code files referenced by the code the States have reviewed that have not been made available for inspection. *Third*, Google's production process has resulted in numerous documents with obfuscated and/or truncated content. For example, Google has produced a number of documents containing mathematical formulas, but key symbols in these formulas have been replaced by nonce characters (*e.g.*, "⬚" or "□"). As another example, a significant number of embedded comments in the documents in Google's production have been only partially produced (*i.e.*, a portion of a comment has been produced, but the rest of the

January 16, 2024
Via E-mail
Page 2

comment has been truncated). Google should immediately reproduce these documents in their complete form, which may require production of the native document. *Fourth*, Google has failed to produce relevant APIs, specifications, programming guides, and onboarding/training materials related to the Display Ads products and systems. *Fifth*, the States are concerned that Google has not fully searched relevant document repositories, including "████████" found at "██████████████" and at "██████."

I. Display Ads Dashboards/Monitoring/Reporting

Numerous Google documents confirm the existence and importance of internal reporting and monitoring tools that Google maintains in the ordinary course of its business, including various onboarding and training documents for Google employees that identify and discuss a variety of monitoring and reporting tools that these employees are expected to use in carrying out their job function. For example, GOOG-AT-MDL-004017319 is a PowerPoint presentation titled ████████████████████" This document explains that RASTA ██ GOOG-AT-MDL-004017339. As another example, GOOG-AT-MDL-000992184 is titled "███████████████" and contains a list of links to various resources, including █████████████████████████ (GOOG-AT-MDL-000992186) and in another document as "█████████████████████████████" (GOOG-AT-MDL-000992438). These two tools, Lumina and RASTA (the Reliable Ads Stats Tracking Architecture) are referenced hundreds of times in the documents Google has produced.

While RASTA and Lumina appear to be ██████████████████████████████, they are far from the only relevant reporting/monitoring tools identified in Google's documents. For example, in addition to referencing the "███████" dashboard, GOOG-AT-MDL-000992438 also discusses ██ And under the heading "Monitoring," GOOG-AT-MDL-000992184 references ██

As the States have noted in prior correspondence, the documents in Google's production frequently include links to other Google documents; based on our current investigation, there are over 300,000 unique links spread across more than 500,000 documents. *See, e.g.*, Sept. 29, 2023 letter from J. Elias to R. McCallum. These links all begin with a "go/…" prefix followed (almost always) by a hyphenated description or title of the document or resource found at the link, e.g., "██████████████████████" or "███████████████████." The States have identified over 2,400 unique "go/…" links that contain the term "dashboard." A list of these

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024
Via E-mail
Page 3

dashboard links and the beginning Bates number of the document(s) in which they are found is attached as Appendix A. The relevance of many of these dashboards is clear from the text of the link alone. To cite just a few examples: [REDACTED]. Given that the links to RASTA and Lumina the States have found do <u>not</u> include the word "dashboard," the list in Appendix A is necessarily incomplete; the States have no way of determining what other, relevant monitoring tools may exist (especially where the term "dashboard" was not used to link to a tool).

Access to these dashboards/monitoring tools and their corresponding reports is necessary given the significant volume of data generated by Google's Display Ads systems. According to a Google document from 2022, Google's "[REDACTED]" (which is described as the "[REDACTED])" *See* GOOG-AT-MDL-001103174-5. The [REDACTED] thus generates more data *daily* than all of the data contained in the Library of Congress. *See* https://www.loc.gov/programs/digital-collections-management/about-this-program/frequently-asked-questions/ (in 2022, the Library of Congress had 21 petabytes of digital content). Given this volume of data, it would be unreasonable and disproportional to expect the States to parse these raw log files, especially when Google has developed tools that it uses to analyze this data. Moreover, the States note that Google has <u>not</u> produced this underlying data, but instead small, incomplete snapshots that are missing numerous data fields. *See, e.g.*, Sept. 29, 2023 letter from D. Pearl to M. Mao; Sept. 14, 2023 letter from D. Pearl to M. Mao; Aug. 25, 2023 letter from D. Pearl to W. Noss; Aug. 15, 2023 letter form D. Pearl to M. Mao; July 27, 2023 letter from W. Noss to D. Pearl; June 26, 2023 letter from M. Mao to D. Pearl (summarizing requests); June 19, 2023 email from M. Mao to D. Pearl; Apr. 10, 2023 email from W. Noss to D. Pearl.

These documents and tools are responsive to at least the States' Requests for Production Nos. 25, 131, 160, 241, 248, 259, 270, and 271, and should be produced (or otherwise made available for inspection) immediately. Courts routinely require parties to produce relevant reports and reporting tools. *See Apple Inc. v. Samsung Electronics Co. Ltd.*, 2013 U.S. Dist. LEXIS 116493, 2013 WL 4426512 , at *3 (N.D. Cal. Aug. 14, 2013) ("Courts regularly require parties to produce reports from dynamic databases, holding that 'the technical burden . . . of creating a new dataset for the instant litigation does not excuse production.'"); *see also McClain v. Lufkin Indus*, 1997 U.S. Dist. LEXIS 25005, at *6-8 (E.D. Tex. Nov. 26, 1997) (requiring production of a database in "machine readable form" to "make more efficient the inevitable discovery"); *see also Humphrey v. Leblanc*, 2021 U.S. Dist. LEXIS 151129 (M.D. La Aug. 11, 2021) (requiring production of data "in a useable form"). Google has agreed to produce such information in at least one other case and has been ordered to produce such information in the past. *See, respectively,*

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024
Via E-mail
Page 4

Brown v. Google LLC, Case No. 20-cv-03664-LHK Dkt. 288 (N.D. Cal. Oct. 1, 2021) (Google agreeing to produce "all internal dashboard information for Incognito mode, Google Analytics, and Google Ad manager) and *Gonzales v. Google, Inc.*, 234 F.R.D. 674, 685, 688 (N.D. Cal. 2006) (denying third party Google's motion to quash government subpoena which required production of information from databases, despite the need for Google to create "new code to format and extract query and URL data from many computer banks"). Google should immediately provide the States with an identification of all relevant dashboards (or other monitoring/reporting tools), any documentation that Google maintains in the ordinary course regarding how to use/operate these dashboards, an identification of the fields and values that can be selected for each dashboard to generate a report, and the date range of data available to be processed by each dashboard. Google should also immediately produce any reports that it generated using these dashboards in the ordinary course of business. Once the States are in possession of this information, Google should then make these dashboards available for inspection in a cleanroom environment so that the States can use these dashboards to generate relevant reporting.

II. Source Code Deficiencies

Google's production of relevant source code is incomplete. The States identify the following deficiencies and request immediate production of all relevant code:

1) <u>Code for the initial Bernanke implementation</u>. It appears that Google has inadvertently omitted the initial Bernanke implementation code. While Bernanke was launched in 2013 and Google has provided a 2013 code snapshot, the snapshot does not contain Bernanke code. Instead, Google's first produced code that contains Bernanke is the snapshot from 2015. Such code is relevant to the DTPA claims regarding Bernanke as it shows how Bernanke worked at launch. Further, such code will permit the States to track and evaluate how Google changed Bernanke over time.

2) <u>Code for Project Bell</u>. Project Bell is identified in the Complaint at ¶¶ 311, 317, and 556-559. The produced Bernanke documents and code further indicate internal options to "execute Project Bell." Google, however, has not produced any code or documents for Project Bell. Such documents and code are relevant and necessary on their face for the asserted claims.

3) <u>Missing files/directories</u>. We have identified a number of relevant source code files/directories that Google has not made available for inspection. For example " ███ ████████████████ " references the following files that have not been made available for inspection: ████████████ ; ███████████████ ; ████████████████████ ; ███████████████████. Such files are relevant to understanding RPO's training and training data, and therefore understanding the operation of RPO.

   Similarly, ████████████ references files in " ███ " that have not been produced, including the main folder and subfolders including " ██████ ", " ███████████ ", and

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024
Via E-mail
Page 5

"███████████████". Other files have similar issues. For example, █████████ references files in the "███████████████" folder that have not been produced. "█████████████" references unproduced files in "████████████████████". "█████████████████████████████████", and "█████████████████████████" reference unproduced code in "███████████" and its subfolders. "█████████████████████████" references unproduced code in ████ ████ █████ ██ ████ ████. "███████████████████" and "███████████████" reference unproduced files in ███████████████████████████████████████.

Furthermore, the folder "████████████████████████" appears to be missing from the June 2023 snapshot. The code in this folder has inputs to the UPR and RPO code that are needed to understand how UPR and RPO operate. Please immediately produce all files in "████████████████████" and all subdirectories.

With respect to Bernanke, Google has not produced the files in "████████████████" or "████████████" that are used by Bernanke, referenced in snapshots, and are relevant to understanding how Bernanke operates and what data is logged.

Please immediately produce all files in the folders and subfolders referenced above, including the following directories and any subdirectories:

- ████
- █████
- ████
- ████████
- ██████████████
- ██████████
- █████████████ (missing from June 2023 snapshot)

4) <u>Missing experiment files</u>. The Bernanke code takes inputs from the experiment code to load values for various bid adjustments. These files appear to typically take the form of SSTable files, which have not been produced. Please additionally produce all experiment files used by the Bernanke code, including but not limited to all of the SSTable files. Examples of such files include, but are not limited to:

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE
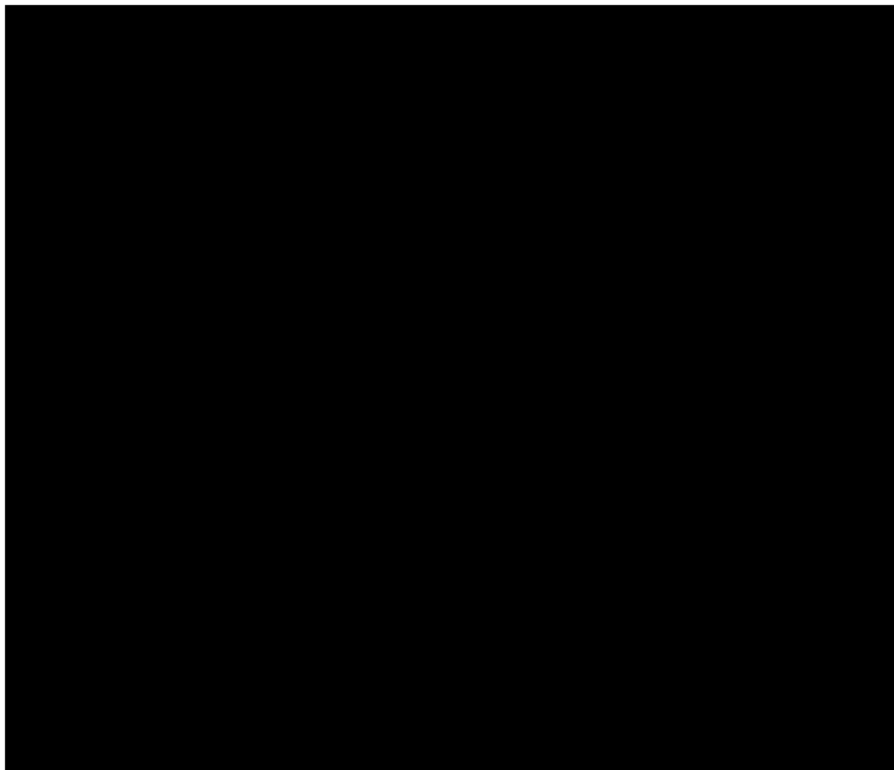
January 16, 2024
Via E-mail
Page 6

Additionally, pursuant to the Confidentiality Order, App'x B at ¶ 10, the States request the Cassandra software be installed on the source code computers to allow its experts to view SSTable files.

III. Obfuscated/Truncated Documents

The States request that Google reproduce documents that have missing, garbled, or otherwise unreadable translations. For example, several Google documents allegedly contain mathematical formulas; however, the produced documents replace key symbols in the formulas with "?" or "□". For example, GOOG-AT-MDL-004334894 contains several pages of formulas, all of which are rendered unreadable:

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024
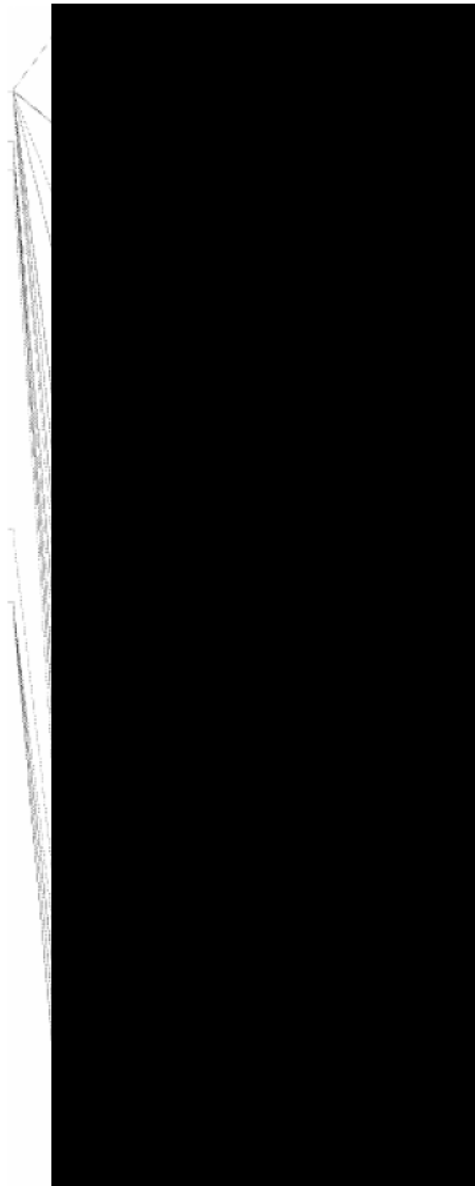Via E-mail
Page 7

Id. at -4906, -4907.

Similarly, many of Google's documents improperly truncate comments. The same -4894 document that contains the unreadable formulas also cuts off comments:

January 16, 2024
Via E-mail
Page 8

Id. at -4897 (cutoff comments designated by ellipses). *See also* GOOG-DOJ-AT-02326571, GOOG-NE-06841582 (containing cutoff comments).

Under the terms of the Order regarding Discovery Procedure, Dkt. 183 (Dec. 13, 2023), all produced TIFFs:

> will show all text and images that are visible in the form in which the electronic document was last saved, with the exception of redacted portions. Hidden content, tracked changes or edits, comments, notes, and other similar information, shall to the extent reasonably practicable, be produced in native or Microsoft-equivalent format, or be imaged so that

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024
Via E-mail
Page 9

such content is viewable on the image file. Where hidden content, tracked changes or edits, comments, notes, or other similar information is not readable in its produced format, a party that received the document may make reasonable requests that a more readable version, such as a native or Microsoft-equivalent version, be produced when helpful to interpret the contents of the relevant document.

Google's document productions fall far short of this standard. Google has produced the vast majority of Microsoft Word documents only in TIFF format, without verifying whether it was truncating comments or maintaining the fidelity of the native. Due to the pervasiveness of these issues, the States request Google reproduce all Microsoft Word documents with corrected TIFFs and in native form. The States further request Google search for and reproduce all documents containing the "question mark in box" or "empty box" characters indicating improper encoding.

IV. Display Ads Document Repositories

The States are concerned that Google has not performed reasonable searches for technical documents. Google's produced documents indicate that ███████████████████ ██ For example, GOOG-NE-07285295 describes ████████████████ ██ ██ Such documents are relevant to the operation of the Display Ads products and systems, the goals of the specific functionalities at issue, and the identification of relevant deponents, yet it does not appear that such directories were searched, nor have such documents been produced.

██ ██ ██ ██ These directories largely do not appear to have been searched nor relevant documents produced. In the rare cases a relevant document has been produced, it appears to be an antiquated version. *See* NRFSUPP2_0008214239 (████████████████ ██ ")█. ██ GOOG-NE-07285295 at -5301, -5302, and that ████████████ ████████████████████. *Id.* at -5303.

Other documents make similar assertions. For example, GOOG-AT-MDL-002462689 requests the recipient "████████████████████████████████████". GOOG-AT-MDL-001397578 refers to ████████████████ ████████████████████████████████ GOOG-AT-MDL-000993528 directs readers to ████████████, while GOOG-AT-MDL-001290493 involves ████████

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024
Via E-mail
Page 10

█████████████████████████. *See also* NRFSUPP_0001885153 ("█████████
█████████"). GOOG-AT-MDL-
000035563 and GOOG-NE-09459207 both state that ████████████████
██████████████████████. GOOG-AT-MDL-003133329 states that
████████████████████████████████████ From the
descriptions in the documents and above, it is plain that these repositories contain a consolidated
set of documents relevant to the operation, use, and goals of the Display Ads products and systems
at issue in the States' DTPA and Antitrust claims.

The States therefore request Google confirm that it has searched and produced all relevant
documents in the locations described above.

V. API and Other Technical Documentation (RFP 259)

RFP 259 requests, among other things, documents related to the operation of the systems
at issue in this case, including at least Bernanke (and its variants, including at least Bell and Global
Bernanke), Reserve Price Optimization ("RPO"), AdMob, DoubleClick for Publishers ("DFP"),
UPR, Header Bidding, Poirot, Elmo, Dynamic Revenue Sharing ("DRS"), Dynamic Allocation
(and its variants), and Jedi (and its variants). Such documents are relevant to the operation of the
products subject to the States' Sherman Act and Deceptive Trade Practices claims.

To date, however, Google has provided barely any relevant technical documents. Google
has failed to produce relevant APIs for Display Ads products and systems, failed to produce
relevant programming guides and design specifications for these products and systems, and failed
to produce relevant onboarding/training materials for these products and systems. Google's own
documents, however, show that these documents exist. For example, GOOG-AT-MDL-
001399058 asking if "███████████████████████," before several
resource pages. Similarly, a "████████████████" (GOOG-AT-MDL-013085982) has
many links for information on pricing rules, UPR, and DRS. These links, however, do not describe
the information provided, and instead merely state that, for example, "additional background <u>here</u>,
<u>here</u>, <u>here</u>, <u>here</u>, <u>here</u>, <u>here</u>, <u>here</u>." *Id.* at 85985 (underlining representing presumed hyperlinks in
the original).

The States therefore request that Google confirm that it will search for and produce all
relevant technical documentation relating to the Display Ads products and systems, including APIs,
specifications, programming guides, and onboarding materials. To the extent Google believes it
has produced all such relevant technical documents for each Display Ads product and system, the
States request Google identify such documents by Bates number.

* * *

Please confirm by **January 23, 2024** that Google will search for and produce (or otherwise
make available for inspection) all relevant documents and code described in the sections above.
Should Google not agree to conduct the requested searches, the States are generally available to

CONTAINS HIGHLY CONFIDENTIAL INFORMATION AND SOURCE CODE

January 16, 2024
Via E-mail
Page 11

meet and confer, and intend to raise any remaining issues with the Judge and/or Special Master. Pursuant to the January 2, 2024 Scheduling Order, the States will further seek to preclude Google from relying on any documents, people, or instruments (including data, reports, or summaries generated using dashboards) that are not promptly disclosed.

Sincerely,

/s/ Zeke DeRose III .
Zeke.DeRose@LanierLawFirm.com
THE LANIER FIRM, P.C.
10940 W. Sam Houston Parkway N.,
Suite 100
Houston, Texas 77064
Telephone: (713) 659-5200
Facsimile: (713) 659-2204

Counsel for Texas, Idaho, Louisiana,
Mississippi, North Dakota,
Mississippi, South Carolina, and South Dakota

Submitted on behalf of all Plaintiff States

Joseph M. Graham, Jr.
joseph.graham@nortonrosefulbright.com
Geraldine Young
geraldine.young@nortonrosefulbright.com
NORTON ROSE FULBRIGHT US LLP
1301 McKinney, Suite 5100
Houston, Texas 77010
(713) 651-5151

Marc B. Collier
Marc.Collier@nortonrosefulbright.com
NORTON ROSE FULBRIGHT US LLP
98 San Jacinto Blvd., Suite 1100
Austin, Texas 78701
(512) 474-5201

Counsel for the Plaintiff State of Texas